

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**Директор физтех-школы
прикладной математики и
информатики**

А.М. Райгородский

	Рабочая программа дисциплины (модуля)
по дисциплине:	Системы хранения данных
по направлению:	Прикладная математика и информатика
профиль подготовки:	А1360: Передовые методы искусственного интеллекта Физтех-школа Прикладной Математики и Информатики кафедра алгоритмов и технологий программирования
курс:	2
квалификация:	бакалавр

Семестр, формы промежуточной аттестации: 4 (весенний) - Дифференцированный зачет

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 30 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 30 час.

Всего часов: 90, всего зач. ед.: 2

Программу составил: О.Н. Ивченко, заведующий кафедрой

Программа обсуждена на заседании кафедры алгоритмов и технологий программирования 12.02.2024

Аннотация

Данный курс призван дать фундаментальные знания в области хранения и обработки данных, для работы с которыми недостаточно одной машины со стандартными аппаратными характеристиками. Будет рассмотрена как батчевая обработка данных, так и обработка в реальном времени.

1. Цели и задачи

Цель дисциплины

- изучить принципы построения и разработки систем хранения и обработки данных;
- получить навыки настройки систем хранения данных, проектирования и разработки процесса наполнения систем хранения данных, реализации запросов к системам хранения данных.

Задачи дисциплины

- приобретение студентами навыков проектирования архитектур, применения специализированных инструментов и разработки программных систем для работы с данными.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-2 Способен использовать современные информационные технологии и программные средства при решении задач профессиональной деятельности, соблюдая требования информационной безопасности	ОПК-2.1 Способен применять современные вычислительную технику и сервисы сети Интернет в области (сфере) профессиональной деятельности
	ОПК-2.2 Знает и умеет применять численные математические методы и прикладное программное обеспечение для решения научных задач в профессиональной области
	ОПК-2.3 Знает основные требования информационной безопасности
ОПК-3 Способен составлять и оформлять научные и (или) технические (технологические, инновационные) отчеты (публикации, проекты)	ОПК-3.2 Владеет на практике методологией составления научно-технических отчетов (проектов)
	ОПК-3.1 Знает основные правила оформления научных публикаций и научно-технической документации, в том числе с использованием прикладного программного обеспечения
	ОПК-3.3 Владеет методами визуального и графического представления результатов научной (научно-технической, инновационной технологической) деятельности в виде отчетов, научных публикаций
ПК-2 Способен самостоятельно или в качестве члена (руководителя) малого коллектива организовывать и проводить научные исследования и их апробацию	ПК-2.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации
	ПК-2.2 Способен планировать и проводить научные исследования самостоятельно или в качестве члена (руководителя) малого научного коллектива
	ПК-2.3 Способен проводить апробацию результатов научно-исследовательской работы посредством публикации научных статей и участия в конференциях

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- основные этапы подготовки данных и основы работы с базами данных;
- существующие методы работы с данными, области их применения, достоинства и недостатки;
- основные цифровые технологии и алгоритмы их работы.

уметь:

- применять технологии отбора, обработки, интерпретации и анализа информации;
- отбирать технологии работы с информацией в зависимости от класса задач в области данных.

владеть:

- методами и технологиями обработки и представления информации;
- технологиями извлечения информации из различных источников.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Теоретические основы баз данных (БД)	5	5		5
2	Проектирование баз данных	5	5		5
3	Детальная разработка баз данных	10	10		10
4	Структурированный язык запросов SQL	10	10		10
Итого часов		30	30		30
Подготовка к экзамену		0 час.			
Общая трудоёмкость		90 час., 2 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 4 (Весенний)

1. Теоретические основы баз данных (БД)

Основные способы хранения БД в ЭВМ. Основы реляционной алгебры. Банки данных и знаний. Типология БД. Документальные БД. Фактографические БД. Гипертекстовые и мультимедийные БД.

2. Проектирование баз данных

Установка, конфигурирование и администрирование систем управления баз данных (СУБД) /Пр/

Структура и функции СУБД. Жизненный цикл БД. Виды установки СУБД (ручная установка, автоматизированная, обновление). Безопасность базы данных. Модель безопасности на основе ролей. Проектирование и разработка БД

3. Детальная разработка баз данных

Аналитическая обработка данных. Базы данных NOSQL. Понятие о BIG DATA. Колоночные базы данных. Основы фракталов. Фрактальная математика. Фрактальные методы в архивации. Управление складами данных. Особенности проектирования и разработки OLTP и OLAP приложений.

4. Структурированный язык запросов SQL

Язык SQL-запросов

Структура и семантика языка запросов

Типы запросов: создание таблиц, заполнение таблиц, выборка, сортировка, отображение полей, вставка, удаление

Типы соединений в запросах

Подзапросы. Агрегирование запросов

Реализация запросов инструментальными средами и средствами

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Необходимое оборудование для лекций: компьютер и мультимедийное оборудование (проектор, звуковая система).

6.Перечень рекомендуемой литературы

Основная литература

1. Компьютерные сети, Computer networks, Электронная версия печатной публикации / . — Санкт-Петербург, Питер, 2021

Дополнительная литература

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Не используются

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

На лекциях используется компьютер и мультимедийное оборудование (проектор, звуковая система),

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студент, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике.

Успешное освоение дисциплины требует:

- посещения студентом всех видов аудиторных занятий;
- ведения конспекта в ходе лекционных занятий;
- качественной самостоятельной подготовки к практическим занятиям, активной работы на них;
- активной самостоятельной и аудиторной работы студента;
- своевременной сдачи преподавателю заданий по аудиторным видам работ.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению: Прикладная математика и информатика
профиль подготовки: АІ360: Передовые методы искусственного интеллекта
Физтех-школа Прикладной Математики и Информатики
кафедра алгоритмов и технологий программирования
курс: 2
квалификация: бакалавр

Семестр, формы промежуточной аттестации: 4 (весенний) - Дифференцированный зачет

Разработчик: О.Н. Ивченко, заведующий кафедрой

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-2 Способен использовать современные информационные технологии и программные средства при решении задач профессиональной деятельности, соблюдая требования информационной безопасности	ОПК-2.1 Способен применять современные вычислительную технику и сервисы сети Интернет в области (сфере) профессиональной деятельности
	ОПК-2.2 Знает и умеет применять численные математические методы и прикладное программное обеспечение для решения научных задач в профессиональной области
	ОПК-2.3 Знает основные требования информационной безопасности
ОПК-3 Способен составлять и оформлять научные и (или) технические (технологические, инновационные) отчеты (публикации, проекты)	ОПК-3.2 Владеет на практике методологией составления научно-технических отчетов (проектов)
	ОПК-3.1 Знает основные правила оформления научных публикаций и научно-технической документации, в том числе с использованием прикладного программного обеспечения
	ОПК-3.3 Владеет методами визуального и графического представления результатов научной (научно-технической, инновационной технологической) деятельности в виде отчетов, научных публикаций
ПК-2 Способен самостоятельно или в качестве члена (руководителя) малого коллектива организовывать и проводить научные исследования и их апробацию	ПК-2.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации
	ПК-2.2 Способен планировать и проводить научные исследования самостоятельно или в качестве члена (руководителя) малого научного коллектива
	ПК-2.3 Способен проводить апробацию результатов научно-исследовательской работы посредством публикации научных статей и участия в конференциях

2. Показатели оценивания компетенций

В результате изучения дисциплины «Системы хранения данных» обучающийся должен:

знать:

- основные этапы подготовки данных и основы работы с базами данных;
- существующие методы работы с данными, области их применения, достоинства и недостатки;
- основные цифровые технологии и алгоритмы их работы.

уметь:

- применять технологии отбора, обработки, интерпретации и анализа информации;
- отбирать технологии работы с информацией в зависимости от класса задач в области данных.

владеть:

- методами и технологиями обработки и представления информации;
- технологиями извлечения информации из различных источников.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

Промежуточный контроль знаний – тестирование:

Раздел 1

1. Модель логического уровня, представляющая собой отображение логических связей между элементами данных безотносительно к среде хранения, называется:

а) инфологической; б) даталогической ; в) физической; г) семантической.

2. Формализованное описание предметной области называется:

а) инфологической; б) даталогической ; в) физической; г) семантической.

3. Модель, определяющая используемые запоминающие устройства и способы физической организации данных в среде хранения называется:

а) инфологической; б) даталогической ; в) физической; г) семантической.

4. Операция Select предназначена для:

а) извлечения определенных строк из таблицы; б) извлечения определенных столбцов из таблицы; в) соединения двух таблиц на основе значений в общих столбцах; г) поддержания целостности ссылок.

5. Операция Join предназначена для:

а) извлечения определенных строк из таблицы; б) извлечения определенных столбцов из таблицы; в) соединения двух таблиц на основе значений в общих столбцах; г) поддержания целостности ссылок.

6. Использование директивы Group by позволяет:

а) отсортировать записи по любому столбцу; б) задания критерия отбора возвращаемых запросов данных; в) сгруппировать все похожие строки результата выполнения запроса по значениям одного или нескольких столбцов; г) объединения нескольких таблиц вместе.

7. Использование директивы Order by позволяет:

а) отсортировать записи по любому столбцу; б) задания критерия отбора возвращаемых запросов данных; в) сгруппировать все похожие строки результата выполнения запроса по значениям одного или нескольких столбцов; г) объединения нескольких таблиц вместе.

8. Использование директивы Where позволяет:

а) отсортировать записи по любому столбцу; б) задания критерия отбора возвращаемых запросов данных; в) сгруппировать все похожие строки результата выполнения запроса по значениям одного или нескольких столбцов; г) объединения нескольких таблиц вместе.

9. Первичный ключ – это:

а) поле или совокупность полей, используемых для установления связей между таблицами; б) поле или совокупность полей обеспечивающих быстрый доступ к записям таблицы; в) поле или совокупность полей однозначно идентифицирующих запись; г) главное поле записи.

10. Кто не является пользователя базы данных:

а) администраторы; б) конечные пользователи; в) прикладные программисты; г) трансляторы.

11. Концептуальный уровень архитектуры:

а) представление информации базы данных такими, какие они есть на самом деле б) обобщенное представление пользователей в) индивидуальное представление пользователя г) представление в памяти

12. В реляционных системах данные рассматриваются пользователем как:

а) таблицы б) объекты и отношения между объектами в) набор древовидных структур г) компьютеризированная системы хранения записей

13. Репликация данных используется для:

а) изменения структуры БД; б) выполнения операций над записями; в) манипулирования объектами БД; г) создания копии БД на рабочей станции.

14. Что такое SQL-сервер:

а) система управления базами данных; б) язык для разработки СУБД; в) язык описания сетевой модели; г) язык манипулирования данными

Примеры контрольных работ:

1. Написать запрос на создание БД, включающей не менее 3 таблиц
2. Написать SQL запрос на выборку данных из заданной БД
3. Написать запрос, включающий сортировку по заданному полю
4. Написать запрос на удаление записей из таблицы

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

1. Архитектура «файл-сервер»
2. Архитектура «клиент-сервер»
3. Архитектура «трехзвенный клиент-сервер»
4. Основные этапы проектирования данных
5. Концептуальная модель данных
6. Представление концептуальной модели моделью БД
7. Сетевая модель данных
8. Иерархическая модель данных
9. Основные понятия реляционной модели данных
10. Многомерная модель данных
11. Элементы реляционной алгебры. Объединение.
12. Элементы реляционной алгебры. Декартово произведение
13. Элементы реляционной алгебры. Пересечение.
14. Элементы реляционной алгебры. Деление.
15. Нормализация БД.
16. Представление данных во внутренней и внешней памяти компьютера
17. Использование индексирования для хранения и выборки данных
18. Использование деревьев индексов
19. Построение связей между таблицами в реляционных БД
20. Типы запросов в реляционных БД.
21. Использование условий и функций в запросах
22. Использование групповых операций в запросах
23. Конструирование многотабличных запросов в БД
24. Конструирование запросов на выборку с помощью языка SQL
25. Фильтрация данных с помощью языка SQL
26. Сортировка данных с помощью языка SQL

Критерии оценивания

Оценка "Отлично" (10) - полностью и вовремя решены все задачи без ошибок. Продемонстрирован грамотный подход к решению задач, реализованы оптимальные алгоритмы, код оформлен в едином удобочитаемом стиле.

Оценка "Отлично" (9) - полностью и вовремя решены все задачи без ошибок. Продемонстрирован грамотный подход к решению задач, реализованы оптимальные алгоритмы.

Оценка "Отлично" (8) - полностью и вовремя решены все задачи без ошибок. Продемонстрирован грамотный подход к решению задач.

Оценка "Хорошо" (7) - полностью решены все задачи. Допущены несущественные ошибки.

Оценка "Хорошо" (6) - полностью решено большинство задач. В некоторых задачах допущены и не исправлены ошибки, либо некоторые задачи решены частично.

Оценка "Хорошо" (5) - полностью решено две трети задач. В некоторых задачах допущены и не исправлены ошибки, либо некоторые задачи решены частично.

Оценка "Удовлетворительно" (4) - полностью решено более половины задач. В остальных задачах допущены и не исправлены ошибки, либо некоторые задачи решены частично.

Оценка "Удовлетворительно" (3) - полностью решено более половины задач.

Оценка "Неудовлетворительно" (2) - решено менее половины задач.

Оценка "Неудовлетворительно" (1) - не решено ни одной задачи.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Дифференцированный зачет может проводиться по итогам текущей успеваемости и сдачи заданий и других видов работ, предусмотренных программой дисциплины и (или) путем организации специального опроса, проводимого в устной и (или) письменной форме.

При проведении устного дифференцированного зачета обучающемуся предоставляется 30 минут на подготовку. Опрос обучающегося не должен превышать одного астрономического часа.

Во время проведения дифференцированного зачета обучающиеся могут пользоваться программой дисциплины, а также справочной литературой, конспектами лекций или другими материалами.